

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

平2-56625

⑮ Int. Cl.⁵

G 06 F 3/06
7/22
12/00
15/16

識別記号

3 0 1 C
3 0 1 R
4 7 0 F
D

庁内整理番号

6711-5B
7313-5B
8944-5B
6745-5B

⑬ 公開 平成2年(1990)2月26日

審査請求 未請求 請求項の数 7 (全17頁)

⑭ 発明の名称 データ共用システムおよびデータ更新方式

⑯ 特 願 昭63-208556

⑰ 出 願 昭63(1988)8月23日

⑱ 発 明 者 中 村 勝 憲 神奈川県小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑲ 発 明 者 倉 野 昭 神奈川県小田原市国府津2880番地 株式会社日立製作所小田原工場内

⑳ 発 明 者 白 柳 芳 朗 神奈川県小田原市国府津2880番地 株式会社日立製作所小田原工場内

㉑ 発 明 者 木 城 茂 神奈川県小田原市国府津2880番地 株式会社日立製作所小田原工場内

㉒ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

㉓ 代 理 人 弁理士 筒井 大和

明 細 書

1. 発明の名称

データ共用システムおよびデータ更新方式

2. 特許請求の範囲

1. 複数の上位処理装置と、これらの上位処理装置に共有されるデータが格納される少なくとも一つの外部記憶装置と、複数の前記上位処理装置と前記外部記憶装置との間に介在し、前記上位処理装置と前記外部記憶装置との間における前記データの授受を制御する記憶制御装置とからなるデータ共用システムにおいて、前記外部記憶装置における固定または可変のデータ格納領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の前記上位処理装置からの任意の前記データ格納領域に対する読み出し操作に際して、当該データ格納領域に対応する前記記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態にする第1のコマンド手段と、当該データ格納領域に対する書込操作

に先立って、前記第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを備えたことを特徴とするデータ共用システム。

2. 前記記憶手段が、当該記憶手段に対応する前記データ格納領域に対する前記上位処理装置からの書込操作に際して加算または減算されるカウンタとして機能する複数のエントリからなり、前記第1のコマンド手段は、前記データ格納領域からの読み出し操作に際して当該データ格納領域に対応する前記エントリの内容を読み出し、前記第2のコマンド手段は、当該データ格納領域に対する書込操作に先立って前記第1のコマンド手段によって読み出された前記エントリの値と現在の値とを比較し、両者の一致・不一致を前記上位処理装置に報知する動作を行うことを特徴とする請求項1記載のデータ共用システム。

3. 前記記憶手段が、複数の前記上位処理装置の各々からの読み出し操作の有無を識別する第1

のフラグと、前記上位処理装置からの書き込み操作の有無を識別する第2のフラグとからなり、前記第1のコマンド手段は、前記データ格納領域に対する読み出し操作に際して当該読み出し操作を実行する前記上位処理装置に対応する前記第1および第2のフラグを特定の状態にし、前記第2のコマンド手段は、当該データ格納領域に対する書き込み操作に先立って、前記第2のフラグの前記特定の状態からの変化の有無を判定して前記上位処理装置に報知する動作を行うことを特徴とする請求項1記載のデータ共用システム。

4. 複数の上位処理装置と、これらの上位処理装置に共有されるデータが格納される少なくとも一つの外部記憶装置と、複数の前記上位処理装置と前記外部記憶装置との間に介在し、前記上位処理装置と前記外部記憶装置との間における前記データの授受を制御する記憶制御装置とからなるデータ共用システムにおいて、前記外部記憶装置における固定または可変のデータ格納

領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の前記上位処理装置からの任意の前記データ格納領域に対する更新処理に伴う読み出し操作に際して、当該データ格納領域に対応する前記記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態にする第1のコマンド手段と、当該更新処理に伴う当該データ格納領域に対する書込操作に先立って、前記第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを備え、前記第2のコマンド手段によって前記記憶手段の内容に変化が認められた場合には、当該データ格納領域に対する前記更新処理のための書込操作を抑止することを特徴とするデータ更新方式。

領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の前記上位処理装置からの任意の前記データ格納領域に対する更新処理に伴う読み出し操作に際して、当該データ格納領域に対応する前記記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態にする第1のコマンド手段と、当該更新処理に伴う当該データ格納領域に対する書込操作に先立って、前記第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを備え、前記第2のコマンド手段によって前記記憶手段の内容に変化が認められた場合には、当該データ格納領域に対する前記更新処理のための書込操作を抑止することを特徴とするデータ更新方式。

5. 前記第2のコマンド手段による判定結果によって前記データ格納領域に対する前記更新処理のための書込操作が抑止された場合には、当該更新操作のやり直しのために、直ちに前記第1

のコマンドの実行および当該データ格納領域からの読み出し動作を実行するようにした請求項4記載のデータ更新方式。

6. 前記記憶手段が、当該記憶手段に対応する前記データ格納領域に対する前記上位処理装置からの書込操作に際して加算または減算されるカウンタとして機能する複数のエントリからなり、前記第1のコマンド手段は、前記データ格納領域からの読み出し操作に際して当該データ格納領域に対応する前記エントリの内容を読み出し、前記第2のコマンド手段は、当該データ格納領域に対する書込操作に先立って前記第1のコマンド手段によって読み出された前記エントリの値と現在の値とを比較し、両者の一致・不一致を前記上位処理装置に報知する動作を行うことを特徴とする請求項4または5記載のデータ更新方式。

7. 前記記憶手段が、複数の前記上位処理装置の各々からの読み出し操作の有無を識別する第1のフラグと、前記上位処理装置の各々からの書

込操作の有無を識別する第2のフラグとからなり、前記第1のコマンド手段は、前記データ格納領域に対する読み出し操作に際して当該読み出し操作を実行する前記上位処理装置に対応する前記第1および第2のフラグを特定の状態にし、前記第2のコマンド手段は、当該データ格納領域に対する書き込み操作に先立って、前記第2のフラグの前記特定の状態からの変化の有無を判定して前記上位処理装置に報知する動作を行うことを特徴とする請求項4または5記載のデータ更新方式。

3. 発明の詳細な説明

本発明は、データ共用システムおよびデータ更新技術に関し、特に、複数の上位処理装置に共用されるディスクファイル装置におけるアクセスの排他制御などに適用して効果のある技術に関する。
〔従来の技術〕

たとえば、複数の中央処理装置と、この中央処理装置に接続される外部記憶装置としてのディスクファイル装置と、この両者間に介在するディス

ク制御装置などからなる汎用の電子計算機システムなどにおいては、個々の中央処理装置がディスクファイル装置に記憶されているデータを更新する処理は、まず目的のデータをディスクファイル装置から主記憶装置に読み出し、次に主記憶装置上のデータの内容を変更し、最後に当該データをディスクファイル装置に書き込むという一連の手順を踏むのが一般的である。

ところで、上述のようなデータ更新において、目的のデータが複数の中央処理装置に共用されるものである場合にはデータの消失が懸念される。

すなわち、複数の中央処理装置が同一データの更新をほとんど同時に開始したとすると、ある中央処理装置が自身の主記憶装置上でデータを変更している間に、別の中央処理装置が前記変更操作が未了のデータをディスクファイル装置から読み出すことが起こり得る。

このようなことが起こると、どちらか後に行われたディスクファイル装置への書き込みによって一方の更新内容が消失してしまうという不都合を

生じる。

〔発明が解決しようとする課題〕

このような問題を回避するために、従来では、たとえばデバイスリザーブコマンドを用いた排他制御技術が実用化されている。

このデバイスリザーブコマンドは主にディスク制御装置が提供するコマンド手段であり、ひとつの中央処理装置が所望の期間だけ特定のディスクファイル装置を占有して、他の中央処理装置が同一ディスクファイル装置をアクセスするのを禁止する働きをするものである。

すなわち、データの更新において、中央処理装置がデータの読み出しに先行してデバイスリザーブコマンドを発行し、更新処理が完了した後（ディスクファイル装置への書き込みが終了した後）、にデバイスリリースコマンドを発行して当該ディスクファイル装置の占有状態を解除することによって、上述の問題を回避することが可能となる。

しかしながら、上述のようなデバイスリザーブコマンドを用いた排他制御には以下のような問題

点がある。

第1の問題点は、排他制御がディスクファイル装置全体に対して行われるということである。

すなわち、一旦デバイスリザーブコマンドが発行されると当該ディスクファイル装置に記憶されている全く別のデータセットのアクセスや、参照のみを目的としたアクセスまでも全面的に禁止してしまうことになり、ディスクファイル装置に対する入出力や処理時間を必要以上に引伸すこととなって、計算機システムの処理能力を低下させる要因となる。

さらに、第2の問題点は、中央処理装置が障害その他不測の要因によって占有状態を解除することができなくなった場合、計算機システム全体が停止してしまうという重大な障害につながる恐れがあるということである。

ところで、更新処理の競合による前述のようなデータ消失の問題は、ディスクファイル装置に記憶されるデータに限らず、たとえば主記憶装置を複数の中央処理装置が共用しているシステム構成

においても同様である。

このような構成における前述の問題の解消手段として、いわゆるコンペア&スワップ命令を備えている中央処理装置が存在する。このコンペア&スワップ命令は、主記憶装置内のデータ更新処理において、中央処理装置が変更したデータを主記憶装置に書き込むときに使用される命令である。

すなわち、コンペア&スワップ命令の動作は、主記憶装置へのデータ書き込みに先立ち、以前に主記憶装置から読み出したデータ（すなわち変更する前のデータ）と現在の主記憶装置のデータを比較し、両者が一致していたなら書き込みを実行するというものである。これにより、特定の中央処理装置がデータを読み出してから書き込むまでの間に他の中央処理装置が同一データを書き替えていたなら、上記の比較において不一致となり、書き込みは行われないので、データ消失の問題は避けられる。不一致の場合、中央処理装置は現在の主記憶装置のデータに対して改めて変更をやり直さなければならないので、当命令はデータを書き

込むかわりに主記憶装置からデータを読み出す動作を行う。

このようなコンペア&スワップ命令の動作は、原理的にはディスクファイル装置のデータ更新にも適用可能であり、上述のデバイスリザーブコマンドを用いた排他制御におけるような問題点を生じることもない。

しかしながら、ディスクファイル装置は記憶媒体が回転しているため、更新対象のデータそのものを比較して更新処理の競合に起因する上述のような不都合を回避しようとするコンペア&スワップ命令の適用はシステムの性能的観点から非現実的である。

なぜなら、変更したデータと現在のデータとの比較を行った後、同一データを書き込む（あるいは読み出す）ためには、ディスクが1回転し、記録・再生ヘッドが当該データに到達するまで待たなければならないため、ディスクファイル装置に対する入出力処理の処理時間を大幅に増大させてしまうという大きな問題を生じるからである。

本発明の前記ならびにその他の目的と新規な特徴は、本明細書の記述および添付図面から明らかになるであろう。

〔課題を解決するための手段〕

本願において開示される発明のうち代表的なものの概要を簡単に説明すれば、以下の通りである。

すなわち、本発明になるデータ共用システムは、複数の上位処理装置と、これらの上位処理装置に共有されるデータが格納される少なくとも一つの外部記憶装置と、複数の上位処理装置と外部記憶装置との間に介在し、上位処理装置と外部記憶装置との間におけるデータの授受を制御する記憶制御装置とからなるデータ共用システムにおいて、外部記憶装置における固定または可変のデータ格納領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の上位処理装置からの任意のデータ格納領域に対する読み出し操作に際して、当該データ格納領域に対応する記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態

なお、データ共用システムに関する同様の技術としては、特公昭54-12300号公報に開示される技術がある。

また、前記コンペア&スワップ命令について記述されている文献としては、株式会社日立製作所製、HITAC Mシリーズ処理装置（M/E Aモード）マニュアル（No 8080-2-083-10）がある。

そこで、本発明の目的は、処理速度の低下を招くことなく、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理の競合に起因するデータの消失を確実に防止することが可能なデータ共用システムを提供することにある。

本発明の他の目的は、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理において、複数の上位処理装置による更新処理の競合に起因するデータの消失を確実に防止することが可能なデータ更新方式を提供することにある。

にする第1のコマンド手段と、当該データ格納領域に対する書込操作に先立って、第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを設けたものである。

また、本発明になるデータ更新方式は、複数の上位処理装置と、これらの上位処理装置に共有されるデータが格納される少なくとも一つの外部記憶装置と、複数の上位処理装置と外部記憶装置との間に介在し、上位処理装置と外部記憶装置との間におけるデータの授受を制御する記憶制御装置とからなるデータ共用システムにおいて、外部記憶装置における固定または可変のデータ格納領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の上位処理装置からの任意のデータ格納領域に対する更新処理に伴う読み出し操作に際して、当該データ格納領域に対応する記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態にする第1のコマンド手段と、当該更新

処理に伴う当該データ格納領域に対する書込操作に先立って、第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを備え、第2のコマンド手段によって記憶手段の内容に変化が認められた場合には、当該データ格納領域に対する更新処理のための書込操作を抑止するようにしたものである。

〔作用〕

上記した本発明のデータ共用システムによれば、たとえば、任意の上位処理装置からの外部記憶装置の任意のデータ格納領域に対する更新処理に際して、外部記憶装置から更新対象のデータを読み出す時点で、第1のコマンド手段を用いて記憶手段の内容を読み出すかまたは当該記憶手段を所定の状態にする段階と、読み出したデータに所定の更新処理を施す段階と、更新後の当該データの外部記憶装置に対する書き込みに先立って、第2のコマンド手段によって記憶手段を参照することにより、前記第1のコマンド手段の実行時から現在までに、当該更新データが格納されていた前記デ

ータ格納領域に対する書き込み操作が行われたか否かを判定する段階と、この第2のコマンド手段の実行により当該データ格納領域に対する書き込みがなかったと判定された場合にのみ、当該データ格納領域に対する更新データの書き込み操作を実行する段階とを経ることにより、複数の上位処理装置が同一のデータ格納領域に対して競合して更新処理を実行することを確実に回避できる。

これにより、たとえば更新の対象となるデータそのものを比較して競合の有無を判定する場合のように、アクセスに対する応答の遅い外部記憶装置に必要以上に頻繁にアクセスすることなく、すなわち処理速度の低下を生じることなく、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理の競合に起因するデータの消失を確実に防止することができる。

また、上記した本発明のデータ更新方式によれば、たとえば、任意の上位処理装置からの外部記憶装置の任意のデータ格納領域に対する更新処理において、更新結果の保存のために更新済のデ

ータを当該データの外部記憶装置における元のデータ格納領域に書き込む操作に際して、当該更新処理中に他の中央処理装置からの当該データ格納領域に対する書込の有無を判断し、他の中央処理装置からの書込があった場合には更新済のデータの書込を放棄する動作を行うので、複数の上位処理装置が同一のデータ格納領域に対して競合して更新処理を実行することを確実に回避できる。

これにより、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理の競合に起因するデータの消失を確実に防止することができる。

〔実施例1〕

以下、本発明の一実施例を用いて説明する。

第1図は本発明の一実施例であるデータ共用システムの構成の一例を示すブロック図であり、第2図、第3図、および第4図は、本実施例における動作の一例を示す流れ図、さらに第5図は、動作中における主記憶装置の内部の状態の一例を示す概念図である。

本実施例では、データ共用システムの一例として汎用の電子計算機システムの場合について説明する。

すなわち、本実施例における電子計算機システムは、複数の中央処理装置1（上位処理装置）と、この複数の中央処理装置1の各々に設けられ、当該中央処理装置1を駆動するプログラムや処理されるデータなどが格納される主記憶装置2と、中央処理装置1が主記憶装置2の一部に作成する後述のチャネル・コマンド・ワード（Channel Command Word：以下、単にCCWまたはコマンドと略記する）に基づいて主記憶装置2と外部との間におけるデータの授受を制御するチャネル装置3（上位処理装置）とを備えており、これらは、前記コマンドやデータさらには各々のステータス情報などが伝達されるインターフェイス線4、5、6を介して相互に接続されている。

さらに、前記チャネル装置3には、ディスク制御装置7（記憶制御装置）を介してディスクファイル装置8が接続されており、ディスク制御装置

7はチャネル装置3からのディスクファイル装置8に対するアクセス要求などのコマンドを解析して、当該チャネル装置3とディスクファイル装置8との間におけるデータの授受を制御する動作を行うものである。

また、チャネル装置3とディスク制御装置7との間、およびディスク制御装置7とディスクファイル装置8との間は、それぞれインターフェイス線9、10によって接続されており、これらの相互間におけるデータ、コマンド、ステータス情報などが伝達されるように構成されている。

ここで、本実施例において、ディスクファイル装置8に格納されているデータに対する更新処理に用いられる一般的なコマンドとしては、たとえばディスクファイル装置8における目的のデータの格納位置を指示するLOCATEコマンド50、目的のデータの読み出しを指示するREADコマンド51、目的のデータの書き込みを指示するWRITEコマンド52、直前のコマンドにおける判定結果による条件分岐を行うTICコマンド5

3などがある。

これらのコマンドの各々は、後続のコマンドを引き続いて実行するか否かを示すコマンドチェインフラグaと、当該コマンドで使用する主記憶装置2上のデータ格納領域のアドレスを示すデータポインタbとを備えている。

また、TICコマンド53のデータポインタbは分岐先アドレスを示している。

そして、これらのコマンドは、主記憶装置2の一部に必要なに応じて後述のようにCCWチェインを構成し、随時、中央処理装置1からの指示によって、チャネル装置3を介してディスク制御装置7に送出されることにより、ディスクファイル装置8と複数の中央処理装置1との間におけるデータの入出力が行われるものである。

この場合、ディスク制御装置7の一部には、前記ディスクファイル装置8におけるデータの所定の格納単位である複数の図示しないトラックの各々に対応するようにアドレスが付与され、各々がカウンタとして機能する複数のエントリ11aか

らなるUPN (UPdate Number) テーブル11を備えており、個々のエントリ11aには各々が対応する図示しないトラックに対して書き込み操作が実行される際にインクリメントされるアップデイト・ナンバUPN₁、UPN₂、...が保持されるように構成されている。

さらに、本実施例の場合には、前述の通常のコマンドの他に、先行するLOCATEコマンド50によって指定された前記トラックに対応する前記UPNテーブル11のエントリ11aを参照し、当該エントリ11aに保持されているアップデイト・ナンバを上位のチャネル装置3の側へ送出する動作を行うRUPN (Read UPdate Number) コマンド54 (第1のコマンド手段) と、予め当該RUPNコマンド54によって読み出されている値と現在のアップデイト・ナンバとを比較し、両者が一致するか否かを判定するCUPN (Compare UPdate Number) コマンド55 (第2のコマンド手段) とを備えている。

また、ディスク制御装置7は、チャネル装置3

から与えられる前述のような一連のコマンドを解読するコマンド解読部12と、このコマンド解読部12からの制御信号12a、12b、12c、...によって起動され、到来する個々の前記RUPNコマンド54、CUPNコマンド55などに基づいて所定の動作を行うRUPN処理部13、CUPN処理部14、さらにはWRITEコマンド52の到来時にディスクファイル装置8に対する書き込み操作を行うと同時に、書き込み動作が成されたトラックに対応するUPNテーブル11のエントリ11aをインクリメントする操作などを行うディスクライト処理部15などを備えている。

また、前記RUPN処理部13、CUPN処理部14、ディスクライト処理部15はバス16を介してUPNテーブル11に接続されており、このバス16を介してUPNテーブル11を参照するものである。

ここで、本実施例におけるデータ更新処理に際して、中央処理装置1が主記憶装置2の一部に作

成するCCWチェーンの一例を第5図に示す。

本実施例では、更新処理に先立って当該更新処理の対象となるデータを中央処理装置1がディスクファイル装置8から読み出すときに使用する読み出しCCWチェーン40は、LOCATEコマンド50と、RUPNコマンド54と、READコマンド51とで構成されている。

また、所定の変更が施された前記データをディスクファイル装置8に書き込む際に使用される書き込みCCWチェーン41は、LOCATEコマンド50aと、CUPNコマンド55と、TICコマンド53、WRITEコマンド52、RUPNコマンド54a、READコマンド51aとで構成されている。

さらに、第5図に示されるように、主記憶装置2の一部には、RUPNコマンド54によって読み出されたUPNテーブル11の所定のエン트리11aの内容が格納されるUPN格納領域42と、READコマンド51、51aによってディスクファイル装置8から読み出されたレコードが格納

されるデータバッファ43とが設けられている。

そして、RUPNコマンド54、CUPNコマンド55およびRUPNコマンド54aのデータポイントbは、同一のRUPN格納領域42を指すように設定され、READコマンド51、WRITEコマンド52、およびREADコマンド51aのデータポイントbは同じデータバッファ43を指すように設定されている。

以下、上述のように構成された本実施例の電子計算機システムにおける更新処理の動作の一例を説明する。

中央処理装置1は、データ更新処理の初めに、読み出しCCWチェーン40を指定して入出力命令を発行する。チャンネル装置3は入出力命令を受取り、ディスク制御装置7およびディスクファイル装置8を選択して、主記憶装置2から最初のコマンドであるLOCATEコマンド50を取出しディスク制御装置7にコマンドとして送る。

ディスク制御装置7のコマンド解読部12は、受け取ったコマンドを解読し、コマンドに応じた

処理部を起動する。ここではLOCATE処理部(図示せず)が起動される。

図示しないLOCATE処理部は、ディスクファイル装置8における目的のデータの格納位置を特定するトラック番号などの情報をインタフェース線9を介して受取り、記憶する。その後、チャンネル装置3に終了ステータスを送る。この終了ステータスはチャンネル終了ビットCHEおよびデバイス終了ビットDVEを含み、これはコマンドの実行が正常に終了したことを表す。

チャンネル装置3はこの終了ステータスを受け取るとLOCATEコマンド50のコマンドチェーンフラグaの値を参照する。

第5図に示す如く、この値は1にセットされており、次のコマンドであるRUPNコマンド54が引き続いて実行されることを指定する。チャンネル装置は、この指定に従い、RUPNコマンド54をコマンド解読部12に送出する。コマンド解読部12は制御信号12aを介してRUPN処理部13を起動する。

このとき、RUPN処理部13は第2図に示す如く動作する。

まずRUPN処理部13は、バス16を介して、UPNテーブル11において前記LOCATEコマンド50で特定されるトラックnに対応するエン트리11aに格納されている現在のアップデイトナンバUPN_nを読み出し、インタフェース線9を介してチャンネル装置3に送り(ステップ201)、チャンネル装置3は、受け取ったアップデイトナンバUPN_nを、RUPNコマンド54のデータポイントbが示すUPN格納領域42にストアする。

その後、RUPN処理部13は、前記のLOCATEコマンド50と同様に、チャンネル終了ビットCHEおよびデバイス終了ビットDVEから成る終了ステータスをチャンネル装置3に送り(ステップ202)、処理を終了する。

先行するRUPN54のコマンドチェーンフラグaが"1"であるので、続いてREADコマンド51が実行される。

この、READコマンド51によって、前記のLOCATEコマンド50によって指定されたディスクファイル装置8のトラックnにおける目的のデータが読み出され、チャンネル装置3は、読み出されたデータを主記憶装置2のデータバッファ43に格納する。

このREADコマンド51の実行が終了すると、チャンネル装置3は、当該READコマンド51のコマンドチェインフラグaが“0”にセットされていることから、読出しCCWチェイン40の終了を検知し、中央処理装置1に入出力処理の終了を報告する。

中央処理装置1は、データバッファ43に格納されたデータに対して所望の変更を行い、その後書き込みCCWチェイン41を指定してチャンネル装置3に入出力命令を発行する。

書き込みCCWチェイン41の最初のコマンドであるLOCATEコマンド50aは、更新後のデータを元の格納位置に格納するため、前記LOCATEコマンド50と同じ、ディスクファイル装

ントリ11aの現在のアップデイトナンバUPN_nを読み出し、読み出したアップデイトナンバUPN_nとチャンネル装置3から受け取った先のアップデイトナンバUPN_nとを比較する(ステップ302)。

ここで、前述のように、任意の中央処理装置1から、ディスクファイル装置8の任意のトラックに対してディスクライト処理部15によって書き込み動作が行われる場合には、UPNテーブル11の当該トラックに対応するエントリ11aの内容がインクリメントされるので、以前の読出しCCWチェイン40を実行したときから現時点(CUPNコマンド55の実行時)までの間に、当該トラックに対して全く書き込み動作が行われていなかったならば、前記の比較結果は一致し、逆に一度でも書き込みが行われていたならば不一致となる。

そして、CUPN処理部14は、前記比較結果が不一致の場合には、終了ステータスとしてチャンネル終了ビットCHEおよびデバイス終了ビット

装置8のトラックnを指定し、この時、チャンネル装置3およびディスク制御装置8は、前述の読出しCCWチェイン40の場合と同様に動作する。

次に、LOCATEコマンド50aのコマンドチェインフラグaが“1”であるので後続のCUPNコマンド55が引き続いて実行され、CUPN処理部14が起動される。

このCUPN処理部14の動作の一例を示すものが第3図である。

同図に示されるように、まず、CUPN処理部14は、前述のRUPNコマンド54の実行時に読み出され、主記憶装置2のUPN格納領域42に格納されているアップデイトナンバUPN_nをチャンネル装置3から受け取る(ステップ301)。すなわち、このアップデイトナンバUPN_nは、READコマンド51で読み出したデータが格納されていた当該トラックnに対応するアップデイトナンバUPN_nである。

次にCUPN処理部14は、RUPN処理部13と同様に動作し、UPNテーブル11の該当エ

DVEをチャンネル装置3に送出し(ステップ303)、一方、前記比較結果が一致した場合には、終了ステータスとしてチャンネル終了ビットCHEおよびデバイス終了ビットDVE、さらには状態修飾ビットSTMをチャンネル装置3に送出する(ステップ304)。

ここで、チャンネル終了ビットCHE、デバイス終了ビットDVEおよび状態修飾ビットSTMを含むステップ304の終了ステータスは、チャンネル装置3にコマンド飛び越しを要求する。

すなわち、チャンネル装置3は、状態修飾ビットSTMを含む前記ステップ304の終了ステータスを受け取ると、直後のTICコマンド53を飛び越してWRITEコマンド52の実行に移る。

WRITEコマンド52は主記憶装置2のデータバッファ43に格納されている所定の変更処理済のデータをディスクファイル装置8の元のトラックnの所定の位置に書き込むことを指示する。

この時、WRITEコマンド52によって起動されるディスクライト処理部15の動作の一例を

第4図に示す。

ディスクライト処理部15は、まず、チャネル装置3から受け取ったデータ、すなわち中央処理装置1によって所望の変更が施されたデータバッファ43の内容をディスクファイル装置8に書き込む(ステップ401)。

次に、UPNテーブル11の該当トラックnのエントリ11aの現在のアップデイトナンバUPN_nを読み出し、その値に1を加算した結果を当該エントリ11aに書き込む(ステップ402)。

その後、ディスクライト処理部15はチャネル終了ビットCHEおよびデバイス終了ビットDVEを含む終了ステータスをチャネル装置3に送る(ステップ403)。

チャネル装置3は、WRITEコマンド56のコマンドチェインフラグaが"0"にセットされていることから、中央処理装置1に一連の書込CCWチェイン41の完了を報告する。

一方、前記CUPN55におけるアップデイトナンバUPN_nの比較結果が不一致であった場合、

対する書き込み動作を放棄して、直ちに、更新処理の再試行に備えるべく、RUPNコマンド54aおよびREADコマンド51aを実行し、当該データが格納されているトラックに対応するカウンタ11aにおける現在のアップデイトナンバUPN_nおよび当該データを再度読み出し、目的のデータに対する一連の入出力処理を終了する。

以上が書込みCCWチェイン41によって実現される動作である。

すなわち書込みCCWチェイン41においては、ディスクファイル装置8の特定のトラックに格納されている特定のデータが、更新のために特定の中央処理装置1によって主記憶装置2に読み出されてから更新結果の保存のための書き込み動作を行うまでの間に、他の中央処理装置1による書き込みが行われていない場合には、当該中央処理装置1による更新処理のための書き込み動作を実行する。

一方、更新のために特定の中央処理装置1によって主記憶装置2に読み出されてから更新結果の

前述したようにCUPN処理部14はチャネル終了ビットCHEおよびデバイス終了ビットDVEのみを含む(すなわち、状態修飾ビットSTMを含まない)終了ステータスを送り、これによってチャネル装置3はCUPNコマンド55に引き続くTICコマンド53を実行する。

このTICコマンド53はチャネル装置3のみに対する指令であり、コマンドとしてディスク制御装置7には送られない。

すなわち、TICコマンド53は、一連のコマンドからなるCCWチェインの実行時における分岐を指令し、本実施例の場合には、分岐先のアドレスを示すデータポイントbはRUPNコマンド54aを指している。

したがって、本実施例の書込みCCWチェイン41においては、CUPNコマンド55におけるアップデイトナンバUPN_nの比較結果が不一致であった場合、WRITEコマンド52を飛び越して、すなわち、中央処理装置1によって更新処理が施されたデータのディスクファイル装置8に

保存のための書き込み動作を行うまでの間に、他の中央処理装置1による書き込みが行われていたならば、他の中央処理装置1の更新結果であるデータの消失を防止すべく現在の当該中央処理装置1による書き込み動作を放棄するとともに、最新の目的のデータに対して所定の変更処理を行うべく、直ちに、目的のデータの再読み出しを行う。

なお、書込みCCWチェイン41が上記のいずれの状態を終了したかを中央処理装置1に認識させるには、たとえばチャネル装置3が入出力処理の終了を中央処理装置1に報告するときに、最後に行ったコマンドの次のアドレスを合わせて報告するという周知の技術によって可能である。

以上説明したように、本実施例によれば、ディスクファイル装置8に格納されているデータを複数の中央処理装置1によって共有し、随時、更新処理などを行うデータ共用システムにおいて、複数の中央処理装置1からディスクファイル装置8の同一のトラック内のデータに対する更新処理が競合した場合でも、特定の中央処理装置1による

更新処理が確実に完了した後に他の中央処理装置 1 による更新処理を行わせることができ、せっかく更新したデータが消失するなどの不都合を未然に防止することができる。

また、複数の中央処理装置 1 の相互間における同一データに対する更新操作の致命的な競合状態の判定を、比較的動作の遅いディスクファイル装置 8 に格納されている実際のデータにアクセスすることなく、高速なアクセスが可能な半導体メモリなどからなる UPN テーブル 11 に設けられた複数のエントリ 11 a の各々の内容の変化に基づいて行うので、競合状態の判定に要する時間が短縮され、更新処理における処理効率が向上する。

さらに、書き込み CCW チェイン 41 において、更新処理の途中に他の中央処理装置 1 によって更新されたデータの消失を防止するために、更新結果の書き込みを放棄する場合には、更新のやり直しを行うべく、直ちに、目的の最新のデータの読み出し処理を実行するので、更新処理の処理効率をより向上させることができる。

さらに第 10 図は、動作中における主記憶装置の内部の状態の一例を示す概念図である。

本実施例においては、前記実施例 1 における UPN テーブル 11 の代わりに、ディスクファイル装置 8 における個々のたとえばトラック単位に対応して設けられた複数のエントリ 21 a を有するフラグテーブル 21 が設けられている。

さらに、この各々のエントリ 21 a には、複数の中央処理装置 A, B, C, . . . の何れからの更新処理が実行中であるかを示す予告フラグ 21 b (第 1 のフラグ) と、当該エントリ 21 a に対応するトラックに対して他の中央処理装置からの書き込みがあったか否かを表示する更新フラグ 21 c (第 2 のフラグ) との対を、複数の中央処理装置 A, B, C, . . . の各々毎に設けられている。

また、コマンドとしては、前記実施例 1 の場合における RUPN コマンドの代わりに、特定のたとえば中央処理装置 A の更新処理における特定のデータの読み出しに際して、当該データが属する

また、競合状態の判定が、ディスクファイル装置 8 におけるトラック単位に行われるので、ディスクファイル装置 8 の全体が特定の中央処理装置 1 によって占有されることがなく、複数の中央処理装置 1 からディスクファイル装置 8 に対するアクセスの待ち時間が大幅に短縮され、複数の中央処理装置 1 とディスクファイル装置 8 との間におけるデータの授受効率が大幅に向上する。

さらに、ディスクファイル装置 8 の全体が特定の中央処理装置 1 によって占有される場合のように、占有状態の解除の失敗によってディスクファイル装置 8 の全体のデータが他のすべての中央処理装置 1 から使用不能になり、データ共用システムの全体の機能が停止してしまうなどの重大な障害が発生することもない。

〔実施例 2〕

第 6 図は、本発明の他の実施例であるデータ共用システムの構成の一例を示すブロック図であり、第 7 図、第 8 図、および第 9 図は、本実施例におけるコマンドによる動作の一例を示す流れ図、さ

トラックに対応するエントリ 21 a における当該中央処理装置 A に対応する予告フラグ 21 b を "1" にすると同時に更新フラグ 21 c を "0" にする動作を行う STFLG (Set Flag) コマンド 56 (第 1 のコマンド手段) が設けられている。

さらに CUPN コマンドの代わりに、更新結果の保存のために当該データのディスクファイル装置 8 の元の格納位置に書き込む動作に先立って、当該中央処理装置 A に対応する更新フラグ 21 c が "1" か "0" かを判定する CHKFLG (Check Flag) コマンド 57 (第 2 のコマンド手段) が設けられている。

一方、これらのコマンドを解析するディスク制御装置 7 a においては、RUPN 処理部および CUPN 処理部に代えて、STFLG 処理部 13 a, CHKFLG 処理部 14 a が設けられている。

さらに、本実施例 2 においては、WRITE コマンド 52 によって動作するディスクライト処理部 15 a は、指示されたデータをディスクファイル装置 8 の所定の格納位置に書き込むと同時に、

当該トラックに対応するエントリ21aにおいて、当該書き込み動作を行おうとする中央処理装置以外の全ての中央処理装置に対応する予告フラグ21bを調べ、当該予告フラグ21bが“1”の場合には、すなわち、同一のトラック内のデータに対する更新中の場合には、対応する更新フラグ21cを“1”にする操作を行うように構成されている。

以下、本実施例2におけるデータ共用システムでの更新処理の一例を説明する。

まず、特定のデータに対して更新処理を行おうとする任意のたとえば中央処理装置Aは、第10図に示されるように、主記憶装置2の一部に読出しCCWチェイン40aおよび書き込みCCWチェイン41aを指定する。

次に、チャネル装置3を介してLOCATEコマンド50をディスク制御装置7aに送出し、ディスクファイル装置8における目的のデータの格納位置(トラック)を指示する。

その後、STFLGコマンド56がディスク制

御装置7aに発行される。

このSTFLGコマンド56によって起動されたSTFLG処理部13aは、第7図に示されるように、まず、当該LOCATEコマンド50によって特定されたディスクファイル装置8の特定のトラックに対応するエントリ21aにおいて、当該中央処理装置Aに対応する位置の予告フラグ21bおよび更新フラグ21cを、それぞれ“1”および“0”にし(ステップ701)、その後、正常に実行が完了したことを示す終了ステータスとして、チャネル終了ビットCHEおよびデバイス終了ビットDVEをチャネル装置3に送出する(ステップ702)。

その後、READコマンド51が引き続いて発行され、ディスク制御装置7aは、予めLOCATEコマンド50によって特定されているディスクファイル装置8の特定トラックから目的のデータを読み出してチャネル装置3に送出し、チャネル装置3は、このデータを主記憶装置2のデータバッファ43に格納し、これにより、読出しCC

Wチェイン40aが完了する。

その後、中央処理装置Aは、主記憶装置2のデータバッファ43に格納されている前記データに対して所定の更新処理を施したのち、書き込みCCWチェイン41aの実行に移る。

まず、LOCATEコマンド50aを発行して、主記憶装置2の一部において更新された前記データのディスクファイル装置8における元の格納位置をディスク制御装置7aに指示する。

その後、CHKFLGコマンド57を発行する。

このCHKFLGコマンド57を受けたディスク制御装置7aは、CHKFLG処理部14aを起動する。

このCHKFLG処理部14aは、第8図に示されるように、これからディスクファイル装置8に書き込もうとするデータの元の格納位置であるトラックに対応するフラグテーブル21のエントリ21aを参照し、当該書き込み動作を行うとする中央処理装置Aに対応する位置の更新フラグ21cの状態が“1”か“0”かを調べる(ステップ8

01)。

そして、当該更新フラグ21cが“1”の場合には、終了ステータスとして、チャネル終了ビットCHEおよびデバイス終了ビットDVEをチャネル装置3に送出する(ステップ802)。

この時、チャネル装置3は、ディスクファイル装置8からの更新処理のためのデータの読出時点から現在までの間に他の中央処理装置B、C、...のいずれからか当該データの元の格納領域であるトラックに対して書込があったものと判断し、この書込データの消失を回避すべく、当該中央処理装置Aによる更新データの保存、すなわちトラックに対する書込動作を放棄するために、TICコマンド53により、WRITEコマンド52を飛ばして、STFLGコマンド56aに実行アドレスを分岐する。

そして、STFLGコマンド56aおよびREADコマンド51aを実行して、直ちに更新処理のやり直しのためのデータの再読出を行う。

また、READコマンド51aのチェインフラ

グaは“0”なので、チャネル装置3は書込みC C Wチェーン41aの実行を完了する。

一方、更新フラグが“0”の場合には、チャネル終了ビットC H Eおよびデバイス終了ビットD V Eとともに状態修飾ビットS T Mをチャネル装置3に送出する(ステップ803)。

この時、チャネル装置3は、更新のためのデータの読出時点から、更新済の当該データのディスクファイル装置8における元のトラックに対する当該書込操作までの間に他の中央処理装置B、C、...からの書込が無かったと判断し、T I Cコマンド53を飛ばしてW R I T Eコマンド52を発行する。

このW R I T Eコマンド52を受けたディスク制御装置7aは、ディスクライト処理部15aを起動する。

このディスクライト処理部15aは、第9図に示されるように、まず、中央処理装置Aによって主記憶装置2のデータバッファ43において更新され、チャネル装置3から送出された更新済のデ

ータを保存すべく、ディスクファイル装置8における元の格納位置(トラック)に書き込む(ステップ901)。

さらに、当該データが書き込まれたトラックに対応するエントリ21aにおいて、当該中央処理装置A以外のすべての中央処理装置に対応する予告フラグ21bを調べ、当該予告フラグ21bが“1”のもの、すなわち同一のデータに対して更新処理を行おうとしているすべての中央処理装置の更新フラグ21cを“1”にする操作を行い、他の同一トラック内のデータに対する更新処理中の中央処理装置に対して、当該トラックへの書き込みがあったことを警告するようにする(ステップ902)。

その後、終了ステータスとして、チャネル終了ビットC H Eおよびデバイス終了ビットD V Eをチャネル装置3に送出し、チャネル装置3はW R I T Eコマンド52のチェーンフラグaが“0”なので、コマンドのチェーンニングを打ち切り、書込みC C Wチェーン41aの実行が完了する。

このように、本実施例2においても前記実施例1の場合と同様に、複数の中央処理装置A、B、C、...に共有されるディスクファイル装置8の同一のトラックなどの記憶領域に格納されたデータに対して複数の中央処理装置がほぼ同時に更新処理を開始したような場合でも、ある中央処理装置による更新が完了したデータによって他の中央処理装置の更新処理が順次開始されるように確実に制御することができ、特定の中央処理装置による更新済のデータが、他の中央処理装置からの更新結果の書込によって、何ら利用されないまま消失してしまうような不都合が確実に防止される。

また、本実施例2の場合には、フラグテーブル21を構成する個々のエントリ21aにおいては、一つの中央処理装置当たり高々2ビットを確保するだけでよく、当該フラグテーブル21を構成する半導体メモリなどの規模を比較的小さくすることができるという効果がある。

以上本発明者によってなされた発明を実施例に基づき具体的に説明したが、本発明は前記実施例

に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

たとえば、上記の各実施例においては、U P Nテーブルおよびフラグテーブルの複数のエントリをディスクファイル装置8の1トラックを単位として対応付けたが、たとえばシリンドラ、レコード、あるいはデータセットを単位とすることも可能である。

その場合、データセットを単位とするためには、ディスク制御装置7においてデータセットを識別する必要があるが、周知のD e f i n e E x t e n dコマンドで当該ディスク制御装置7に送られるエクステンション情報(データセットの始まりと終わりのトラックアドレス)を利用することで可能となる。

また、U P Nテーブルやフラグテーブルは実在する全てのトラック(記憶単位)の各々に対応するエントリを格納可能な大きさであることが望ましいが、必ずしもそうでなくてもよい。

すなわち、たとえば上述の記憶単位を1トラックとした場合、UPNテーブル11が格納可能なエントリの数をディスクファイル装置のトラックの総数より少なくして構成することができる。

上述の如き構成は、RUPN処理部、CUPN処理部、STFLG処理部、CHKFLG処理部およびディスクライト処理部において、現在更新対象となるデータ記憶領域（トラック）に対応するエントリを参照あるいは更新するとき、該当エントリがUPNテーブルまたはフラグテーブルに登録されているか否かを検索すること、および検索の結果、該当エントリが登録されていなかったときには、UPNテーブルまたはフラグテーブルの他のエントリを削除して新しいエントリに置換することによって可能である。

この場合の検索および置換技術には、たとえばハッシング検索技術やLRU (Least Recently Used) 置換技術あるいはFIFO (First In-First Out) 置換技術など、いずれも周知の技術を用いることができる。

たとえば、任意の上位処理装置からの外部記憶装置の任意のデータ格納領域に対する更新処理に際して、外部記憶装置から更新対象のデータを読み出す時点で、第1のコマンド手段を用いて記憶手段の内容を読み出すかまたは当該記憶手段を所定の状態にする段階と、読み出したデータに所定の更新処理を施す段階と、更新後の当該データの外部記憶装置に対する書き込みに先立って、第2のコマンド手段によって記憶手段を参照することにより、前記第1のコマンド手段の実行時から現在までに、当該更新データが格納されていた前記データ格納領域に対する書き込み操作が行われたか否かを判定する段階と、この第2のコマンド手段の実行により当該データ格納領域に対する書き込みがなかったと判定された場合にのみ、当該データ格納領域に対する更新データの書き込み操作を実行する段階とを繰ることにより、複数の上位処理装置が同一のデータ格納領域に対して競合して更新処理を実行することを確実に回避できる。

これにより、たとえば更新の対象となるデータ

〔発明の効果〕

本発明によるデータ共用システムによれば、複数の上位処理装置と、これらの上位処理装置に共有されるデータが格納される少なくとも一つの外部記憶装置と、複数の前記上位処理装置と前記外部記憶装置との間に介在し、前記上位処理装置と前記外部記憶装置との間における前記データの授受を制御する記憶制御装置とからなるデータ共用システムにおいて、前記外部記憶装置における固定または可変のデータ格納領域の各々に対応して設けられ、当該データ格納領域に対する書込操作の実行によって内容が変化する記憶手段と、任意の前記上位処理装置からの任意の前記データ格納領域に対する読み出し操作に際して、当該データ格納領域に対応する前記記憶手段の状態を読み出すかまたは当該記憶手段を特定の状態にする第1のコマンド手段と、当該データ格納領域に対する書込操作に先立って、前記第1のコマンド手段の実行時からの当該記憶手段の内容の変化の有無を判定する第2のコマンド手段とを備えているので、

そのものを比較して競合の有無を判定する場合のように、アクセスに対する応答の遅い外部記憶装置に必要以上に頻りにアクセスすることなく、すなわち処理速度の低下を生じることなく、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理の競合に起因するデータの消失を確実に防止することができる。

また、上記した本発明のデータ更新方式によれば、たとえば、任意の上位処理装置からの外部記憶装置の任意のデータ格納領域に対する更新処理において、更新結果の保存のために更新済のデータを当該データの外部記憶装置における元のデータ格納領域に書き込む操作に際して、当該更新処理中に他の中央処理装置からの当該データ格納領域に対する書込の有無が確実に判断され、他の中央処理装置からの書込があった場合には更新済のデータの書込を放棄する動作を行うので、複数の上位処理装置が同一のデータ格納領域に対して競合して更新処理を実行することを確実に回避できる。

これにより、複数の上位処理装置に共有される外部記憶装置に格納されたデータに対する更新処理の競合に起因するデータの消失を確実に防止することができる。

4. 図面の簡単な説明

第1図は本発明の一実施例であるデータ共用システムの構成の一例を示すブロック図、

第2図は本実施例における動作の一例を示す流れ図、

第3図は同じく本実施例における動作の一例を示す流れ図、

第4図は同じく本実施例における動作の一例を示す流れ図、

第5図は動作中における主記憶装置の内部の状態の一例を示す概念図、

第6図は本発明の他の実施例であるデータ共用システムの構成の一例を示すブロック図、

第7図は本実施例におけるコマンドによる動作の一例を示す流れ図、

第8図は同じく本実施例におけるコマンドによ

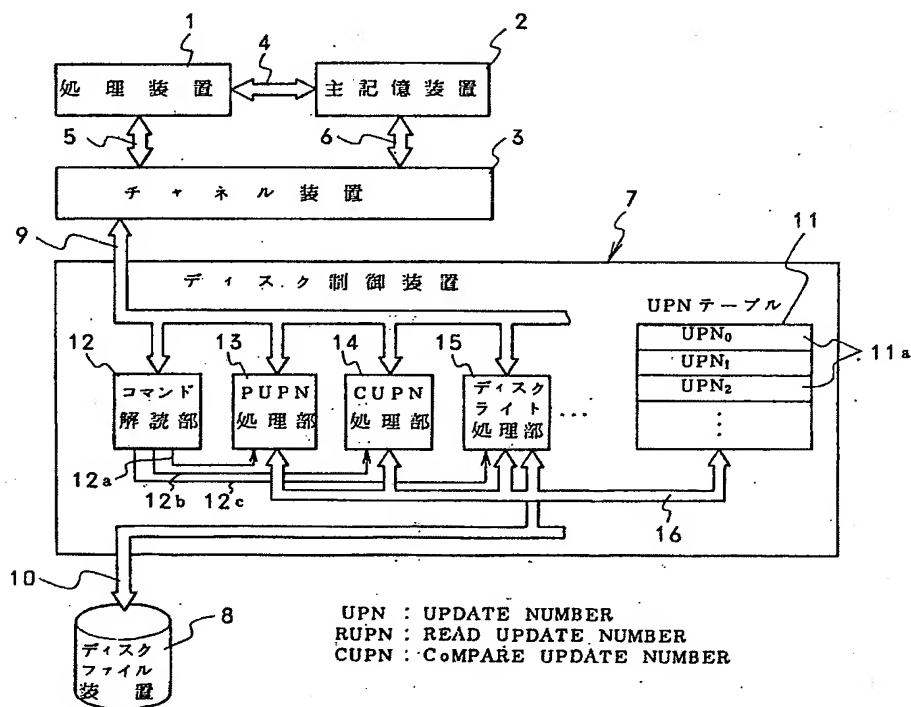
る動作の一例を示す流れ図、

第9図は同じく本実施例におけるコマンドによる動作の一例を示す流れ図、

第10図は動作中における主記憶装置の内部の状態の一例を示す概念図である。

1, A, B, C, . . . 中央処理装置、2 . . . 主記憶装置、3 . . . チャンネル装置、4, 5, 6 . . . インターフェイス線、7, 7a . . . ディスク制御装置、8 . . . ディスクファイル装置、9, 10 . . . インターフェイス線、11 . . . UPNテーブル(記憶手段)、11a . . . エントリ、12 . . . コマンド解読部、12a, 12b, 12c . . . 制御信号、13 . . . RUPN処理部、13a . . . STFLG処理部、14 . . . CUPN処理部、14a . . . CHKFLG処理部、15, 15a . . . ディスクライト処理部、16 . . . バス、21 . . . フラグテーブル、21a . . . エントリ、21b . . . 予告フラグ、21c . . . 更新フラグ、42 . . . UPN格納領域、43 . . . データバッファ。

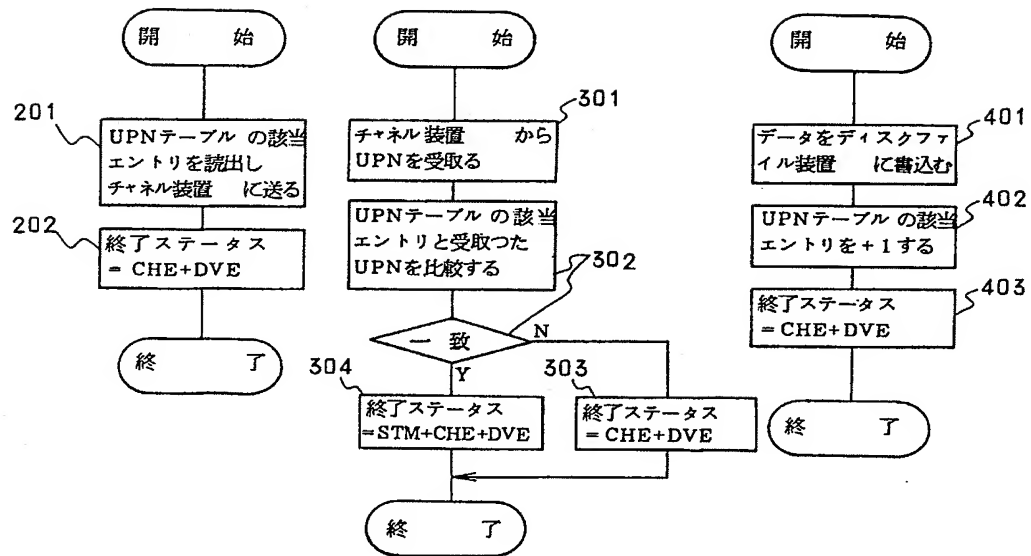
第1図



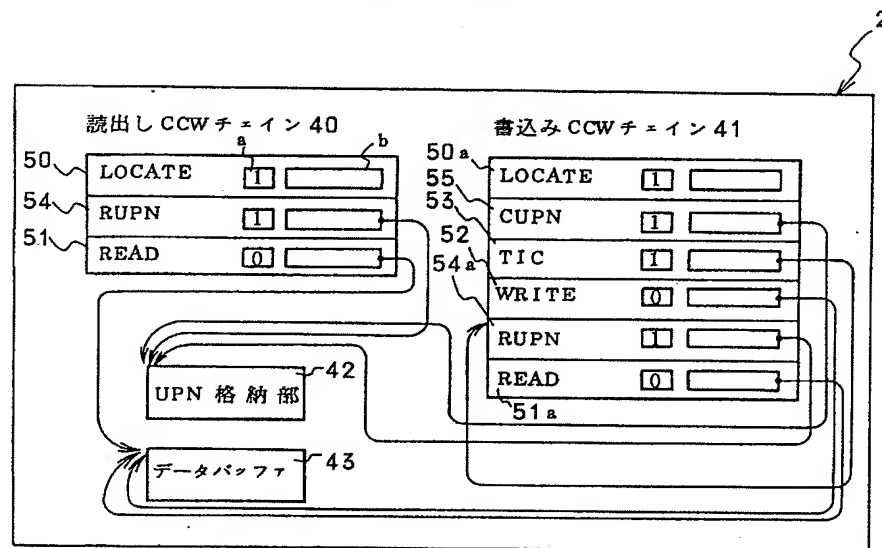
第 2 図

第 3 図

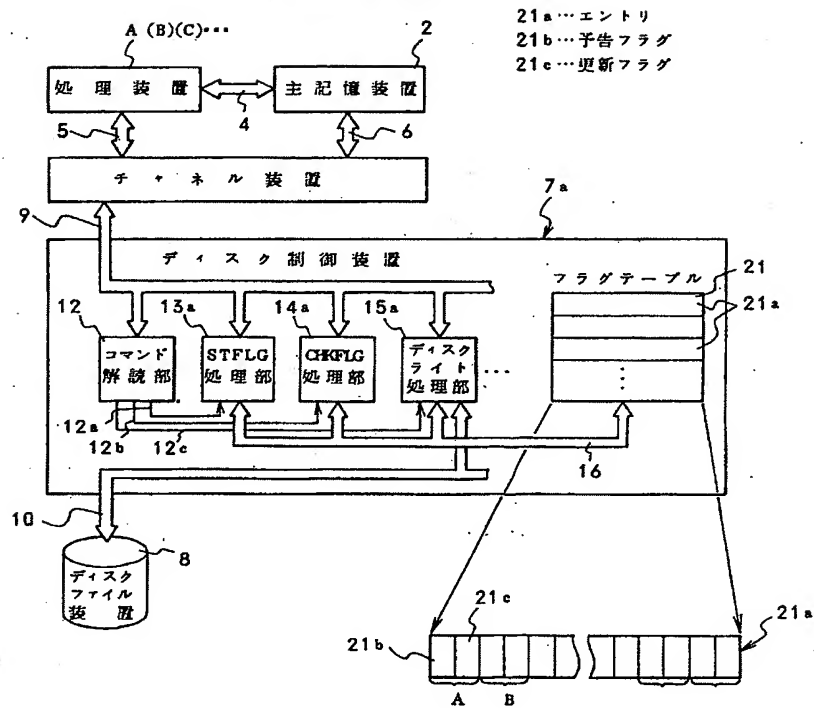
第 4 図



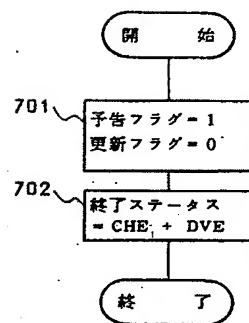
第 5 図



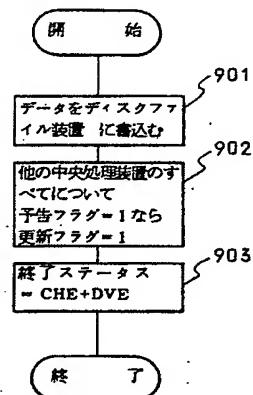
第 6 図



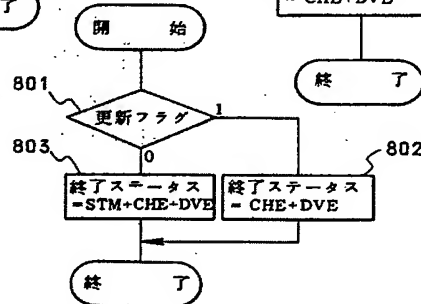
第 7 図



第 9 図



第 8 図



第 10 図

